Demonstrations: Clues to Effective Animated Explanations?

Sandra C. Lozano and Barbara Tversky

Stanford University, Psychology Department, Stanford, CA 94305-2130 USA
scl@psych.stanford.edu; bt@psych.stanford.edu

Abstract. In accordance with the principle of congruence, animation should be a particularly effective tool for conveying information about physical and metaphorical changes over time. However, in most instructional situations, animated graphics are no more effective than static ones for achieving these ends, sometimes even impairing performance. The present research uses people's demonstrations to provide a foundation for enhancing animations. Participants assembled a TV cart and then made a video showing others how to do it, using either gestures and language or gestures alone; a control group merely reassembled the object. Adding communication to the assembly task altered the manner and efficiency with which participants performed it. Demonstrators segmented the assembly into steps, made actions visible to the camera, and added pointing and exhibitive gestures to indicate parts and how to assemble them. These devices were frequent in the communicative conditions, especially the condition without language, indicating that language and gesture supplement as well as complement each other. People's demonstrations suggest ways to enhance animations: break them into steps; use extra-pictorial devices such as arrows (points) and brackets (exhibits), and accommodate viewer's perspective.

1 Introduction

In principle, animated graphics should be particularly effective for conveying processes of change, real and metaphorical. In practice, animated graphics are no more effective than static ones for conveying concepts from the physical to the biological to the abstract [1-3]. One reason animations so often fail as effective learning tool is perceptual: animations are complex and fleeting, whereas static graphics can show successive steps of a complex action sequence simultaneously so they can be inspected and compared [3]. Another reason is conceptual: people frequently conceive continuous events to be sequences of discrete steps [4]. Animations are typically temporally uniform enactments, so they do not break down processes into steps [3]. Effective static diagrams do not reflect space uniformly; maps, for example, enlarge features such as roads and rivers, that would not be visible at map scale. Maps select information as well as distorting it; they present only the relevant information, and schematize that information to highlight the essential details [5-6]. The belief that animation should be an optimal tool for learning is predicated on an assumption that animation enhances comprehension and memory, and that animation promotes insight and inference. Proponents of the idea that animations promote learning also argue that animations attract attention, focus thoughts, and increase the motivation to learn [2].

For any given graphic display to be effective, at a minimum it should adhere to two key principles. First effective visualizations should meet the Congruence Principle, which states that the content and format of the animation must correspond to the desired content and format of the concepts being conveyed. Second effective visualizations should satisfy the Apprehension Principle, which states that the animation should be readily and accurately perceived and comprehended [3]. Presumably then, animation should be particularly well-suited for conveying real-time online processes as well as other information that is inherently dynamic or temporal.

Given the lack of success of most instructional animations, the challenge is to find ways to augment animations so that they can be effective teaching tools. The fact that most animations are continuous, real-time, and unembellished is puzzling, given the sophistication of static graphics. One way that static diagrams are made more effective is by highlighting essential information. Another is by the addition of extra-pictorial devices, notably arrows, guidelines, boxes, and brackets. These devices have context-dependent meanings that are readily apparent to observers [7]. Many of these extra-pictorial devices appear to have parallels in gesture: arrows and pointing, for example, or brackets and showing.

Effective communication evolves from cycles of production and comprehension, where communities revise and refine linguistic devices to promote mutual understanding [8]. One way to short-cut that process is to have participants create communications, then test them for understanding [9-11]. Studying how people demonstrate accomplishing a task, then, should give clues to designing more effective animations for instructing others in accomplishing the task. The demonstrators are expected to segment the continuous action into natural steps; they are also expected to add demonstrative gestures that aid apprehension and congruence

There have been heated interchanges as to the role of gestures in communication. Some argue that they primarily function to promote the speaker's reasoning, and there is support for that [12]. Preventing gesture while describing spatial relations increases disfluencies in speech [13]. Duration of representational gestures increase with increased asynchrony between gesture and lexical affiliate onset, suggesting that gestures serve to maintain conceptual features in memory during lexical search [14].

Others propose that gestures are meant to communicate to others, and are adapted to those ends; there is support for this view as well [15]. According to this view, gestures and their lexical affiliates are coordinated temporally and these two behaviors share a common origin, namely communicative intention [16]. Consistent with this view, interpretations of gestures are context-bound. When participants assign gestures to semantic categories, their assignments of meaning are heavily influenced by language, despite instructions to attend to gesture alone [17]. This parallels assignment of meaning to extra-pictorial devices in diagrams; their meanings are context-bound as well [7].

The present project has both theoretical and applied aims. One theoretical aim is to achieve a better understanding of the role of gestures and gestures and speech in demonstrations. Another is to reveal people's understanding of continuous events, in this case, an assembly task. The applied aim is to reveal ways to augment animated diagrams to increase their effectiveness.

In the present study, people first learned to assemble a TV cart by using the photograph of a fully built one as a guide; then they made a video in which they reassembled it. Participants in the noncommunicative control condition reassembled an identical TV cart; those in the gesture and speech condition showed others how to assemble the TV cart; and those in a gesture with no speech condition showed others who may not understand English how to assemble the TV cart. Assembly is representative not only of a large number of tasks people enact in their daily lives, but also of the behavior of systems, such as scientific or political processes. Since spatial ability has been related to success of both assembly of the TV cart and to effectiveness of instructional design, spatial ability will be assessed and related to performance. The immediate expectation is that gestures will increase from the control to the gesture-speech to the gesture alone conditions, especially gestures of pointing and exhibiting. Also, demonstrators are expected to accommodate a viewer's perspective; that is, to make sure that assembly actions are visible to the camera. In addition, success in assembly and demonstration is expected to correlate with spatial ability.

2 Method

2.1 Participants

Thirty-seven Stanford undergraduates participated for course credit. Data from one participant were excluded due to failure to follow instructions. The results reported are based on the data of 19 males and 17 females.

2.2 Materials

Materials consisted of a paper-and-pencil test measuring spatial ability, questionnaires and two identical television carts. The television carts are 17"x 25"x 21" in size; they consist of two side boards, a lower shelf, an upper shelf, a board to stabilize the upper shelf, pegs for attaching the stabilizer board, screws, screwdriver, and wheels.

2.3 Design

Each participant was randomly assigned to one of three conditions for the second assembly of the TV cart: Gesture, Speech and Gesture, or Control. Participants in the two communication conditions, Gesture Only, and Gesture and Speech, received instructions as follows: "Many people believe that they can learn a novel task best when they see someone else show them how to do it. Now that you are knowledgeable about how to assemble a TV cart, we would like you to please make a videotape in which you clearly demonstrate to someone else how to do this task." The instructions then diverged on the basis of a whether a participant was in the Gesture Only or the Gesture and Speech condition.

In the Gesture Only condition, speech was prohibited; these participants were instructed that "Since your demonstration might be viewed by a non-English speaker, the videotape will not have a soundtrack. You may use gestures and any strategies other than speech that you think will best communicate cross-linguistically how to assemble the TV cart." In contrast participants in the Gesture and Speech condition received instructions as follows: You may use speech, gestures, and any other strategies that you think will best communicate how to assemble the TV cart."

Unlike participants in the communicative conditions, Control participants received the following instructions: "We are interested in how learning and experience can improve performance on an assembly task. Now that you have had practice building a TV cart and are knowledgeable about the task, we would like you to now assemble a new, identical TV cart."

2.4 Procedure

First, participants completed the Vandenburg Mental Rotation Test of spatial ability [18]. This test correlates with behavior in a broad array of spatial tasks, including assembly of the TV cart and quality of instructions produced in previous work [19]. Next all participants assembled a TV cart using the photograph on the carton (Figure 1) and all the necessary parts and tools. Finally, participants assembled a second TV cart according to one of three conditions, described above. The photograph of the TV cart was removed during reassembly. The experimenter was not present during either assembly or reassembly, so as to prevent the occurrence of demand characteristics.

Figure 1 Illustration of the television cart. The figure shows the photograph of the TV cart that participants studied when initially learning the assembly task.

3 Results

Do demonstrators make assembly actions visible to the camera? To answer this, videos from the second assembly were coded for whether or not each step of attachment (e.g., attaching the left sideboard to the left side of the top shelf) was visible to the camera as well as the frequency with which other assembly-relevant gestures (e.g., tightening screws) were visible to the camera. An alpha level of 0.05 was used for all statistical tests reported. An ANOVA was conducted to compare the frequency with which assembly-relevant actions and action steps were made visible to the camera as a function of communication level. More assembly steps were made visible to the camera when people demonstrated the task rather than merely reperformed it, $F(2, 33)=97.94$, $p<0.001$ (Figure 2). Post hoc tests (Tukey's) revealed that participants in the communicative conditions were more likely to make their actions visible to the camera (gestures only $M=8.83$, $SEM=0.30$ and gestures and speech $M=5.92$, $SEM=0.58$) than participants in the control condition ($M=0.67$, $SEM=0.31$). Demonstrators showed assembly to the camera far more frequently than reassemblers, an effect strengthened by the prohibition on speech. For the gesture only group, nearly all the assembly steps were made visible to the camera, on average, 8.83 of the 10 assembly steps.

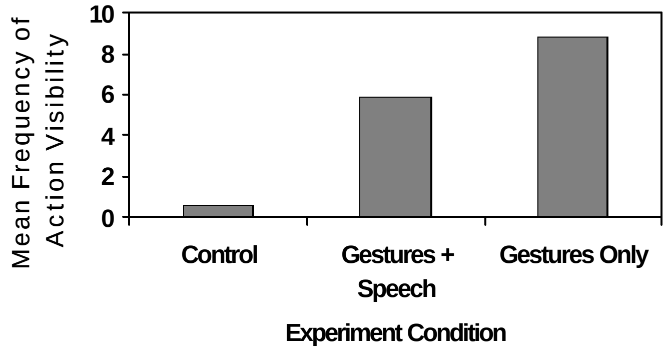

Figure 2 Mean total frequency of actions visible to a viewer as a function of whether and how the assembly task is communicated.

When do demonstrators point and when do they exhibit? Two types of gesture were coded, exhibiting and pointing. Exhibiting occurred when someone held up a piece for the camera to see, and pointing occurred when someone used a finger to point or trace a piece sitting on the table or already in hands. The first author coded all tapes twice for every instance of pointing or exhibiting every component, checked for reliability, and in 5.1% of cases where there was disagreement, came to a resolution. Both exhibiting and pointing are ways of indicating. As is evident in Figures 3 & 4, pointing and exhibiting were frequent in the communication conditions, especially for the condition that did not use language; these gestures did not occur in the control condition. Greater use of exhibitive and pointing gestures in communicative conditions was confirmed by a one-way ANOVA. More total gestures were used in communicative conditions, $F = 37.41$, $p<0.01$. Post hoc comparison testing revealed that participants used significantly more exhibitive gestures when the task was communicative (M=15.08, SEM=2.07) than when it was noncommunicative (M=0, SEM=0). More exhibitive gestures occurred in the gestures only condition than in the control (M=10.33, SEM=0.70); and more exhibitive gestures occurred when communicating without language than with it.

Next, pointing and exhibiting gestures were divided into those referencing small connectors like screws and screw holes and those referencing large components of the TV cart like the top shelf. More gestures were directed toward large object parts in the communicative conditions than in the control condition, $F(2, 33) =37.27$, $p<0.01$ (Figure 3). Post hoc testing revealed that participants in the control condition used significantly fewer exhibitive gestures to refer to large object parts (M=0, SEM=0) than both participants in the communicative condition with gestures only (M=8.42, SEM=1.15) and the communicative condition with speech and gestures combined (M=6.67, SEM=0.51). The two communicative conditions did not differ significantly in mean total number of exhibitive gestures directed at large parts, $p=0.22$.

Regarding small connectors, as seen in Figure 3, participants in the control condition used significantly fewer exhibitive gestures to refer to connectors (M=0, SEM=0) than both participants in the communicative condition of gestures only (M=6.67, SEM=3.27) and the communicative condition of speech and gestures combined (M=3.67, SEM=2.02), $F=18.00$, $p<0.01$. Further post hoc comparison testing showed that participants used significantly more exhibitive gestures to refer to small object connectors when speech was unavailable than when it was available; when speech was available many gestural references to object wholes were omitted or suppressed.

Participants in the control condition (M=0, SEM=0) as well as participants in the communicative condition with gestures and speech (M=3.42, SEM=1.25) used significantly fewer pointing gestures than participants in the communicative condition with gestures only (M=12.83, SEM=3.62), $F(2, 33) =9.03$, $p<0.01$ (Figure 4). Post hoc tests confirmed that participants communicating through speech and gestures were as likely to fail to use pointing gestures during reassembly as participants in the control condition. As with the analysis of the exhibitive gestures, all of the pointing gestures were categorically separated on the basis of whether they referenced large object parts or small object connectors. Participants in the control condition (M=0, SEM=0) as well as participants in the communicative condition of gestures and speech (M=0.17, SEM=0.30) used significantly fewer pointing gestures to refer to large object parts than participants in the communicative condition of gestures only (M=2.08, SEM=1.70), $F(2, 33) =14.51$, $p<0.01$. Likewise participants communicating through gestures only pointed to a greater total number of object connectors (M=10.75, SEM=3.33) than participants communicating through both speech and gestures (M=3.25, SEM=1.19). In neither the case of pointing to object parts ($p>0.05$) nor pointing to object connectors ($p>0.05$) did a significant interaction occur between the control condition and the communicative condition in which speech was available. Participants in these two conditions were equally unlikely to employ pointing gestures to refer to either object parts or object connectors.

Overall, when some form of communication was present, participants used significantly more total exhibitive gestures (M=12.71) than total pointing gestures (M=8.13), t(22)=2.94 (Figure 5). Similarly, for referring to object parts, participants used significantly more exhibitive gestures (M=7.54) than pointing gestures (M=1.15), t(22)=10.12. However, exactly the opposite pattern was found for referencing object connectors; participants used significantly more pointing gestures (M=7.00) than exhibitive gestures (M=5.17) to refer to object connectors, t(22)=1.25.

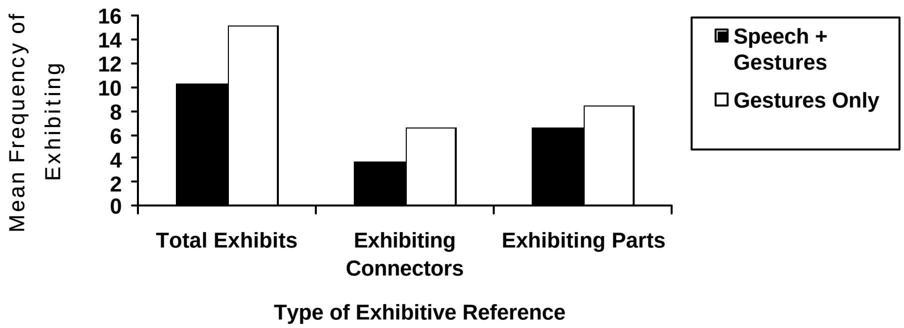

Figure 3. Mean frequency of exhibiting gestures by condition and function. No exhibitive gestures occurred in the control condition; they frequently occurred, though, in the two communicative conditions.

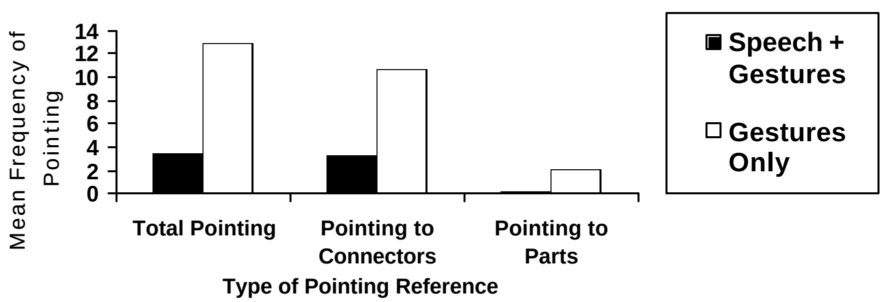

Figure 4 Mean frequency of pointing gestures to by condition and function. No pointing gestures occurred in the control condition; they frequently occurred, though, in the two communicative conditions.

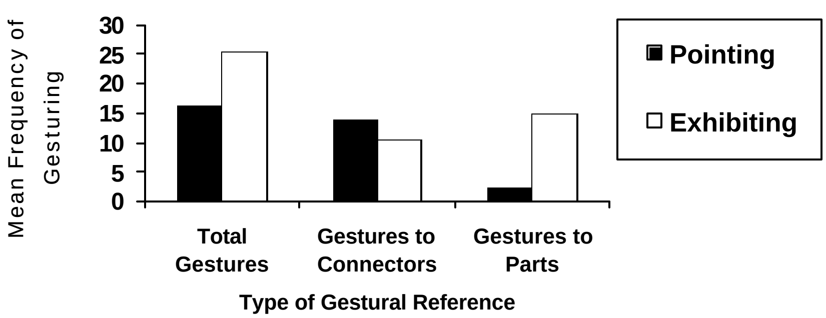

Figure 5 Differences in mean frequency of pointing vs. exhibiting of TV cart components.

How do demonstrators segment actions into meaningful steps? Participants in the two communicative conditions frequently made efforts to segment the reassembly process into discrete, meaningful steps. In contrast, participants in the control condition never engaged in activities of this sort. A one-way ANOVA confirmed this greater usage of step markers when reassembly was communicative than when it was noncommunicative, $F(2, 33) = 84.60$, $p<0.01$ (see Figure 6). For participants in the speech and gesture condition, the most common way of segmenting assembly into discrete steps was to use verbal time markers to signal step initiation (e.g., "The final step is to attach the wheels."). The other common, though less frequently used, technique was to use verbal time markers to signal step completion (e.g., "Now, that these 2 boards are attached, this part of the cart is complete."). Participants in the speech and gesture condition never used gestures to signal either step initiation or step completion; actions were segmented purely through the verbal channel. Although participants in the speech and gesture condition never used gestures for segmenting actions into steps, participants in the gesture only condition often used gestures to communicate information about which actions grouped into meaningful steps. Gestures used by participants in the gesture only condition were coded for whether or not they conveyed information about step segmentation and then for whether they conveyed information about step initiation (e.g., holding up two fingers to indicate starting the second step of assembly) or step completion (e.g., using hand to make an "okay" sign or a "thumbs up" sign once a step is complete). This type of gestural behavior could then be compared to the verbal behavior of participants in the speech and gesture condition as well as to the behavior of control condition participants. As evidenced by the frequency with which they used gestures signaling step initiations and completions, participants in the gesture only condition did frequently make efforts to segment actions into discrete steps (M=3.33, SEM=0.57), though to a significantly lesser extent than did the speakers (M=7.83, SEM=0.47), as evidenced by post hoc comparison testing.

Both for signaling step initiations, $F(2, 33) = 62.93$, $p<0.01$, and for signaling step completions, $F(2, 33) = 12.48$, $p<0.01$, participants in the communicative conditions clearly segmented more actions on these bases than did control condition participants. There were 10 steps altogether, so participants in the speech and gesture condition marked 78.33% of the steps on average and participants in the gesture only condition marked 38.33% of the steps on average. Whereas speakers marked step initiations significantly more than they marked step completions $t(22)=5.52$, $p<0.05$, gesture-only participants marked step completions significantly more than they marked step initiations, $t(22)=1.99$, $p<0.05$. These differences between conditions likely reflect the fact that it is easier to convey information about time and step ordering verbally than gesturally.

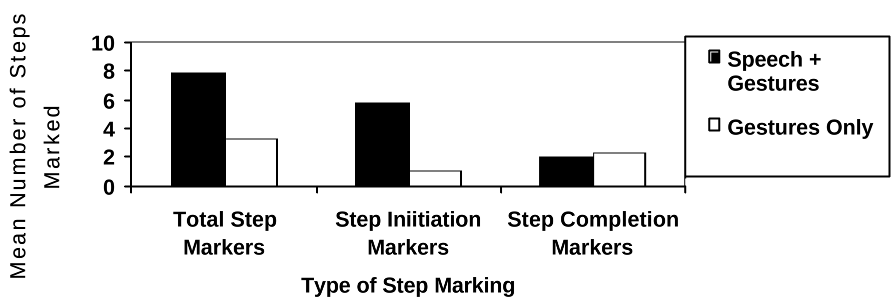

Figure 6 Differences in mean number of actions segmented into steps

How do demonstrators express assembly? Speech, signs, and gestures all serve as natural means for describing and experiencing environments, both from within and from without [20]. Another way demonstrators segmented action into meaningful units was by constructing models with their hands and arms. Prior to actually performing a step, demonstrators often created gesture models either showing the actions that should be performed for each step or showing what the object structure would look like upon completion of that step. A string of gestures was coded as a model when three or more successive gestures coordinated to portray structure and/or action. As confirmed by a one-way ANOVA, demonstrators produced significantly more gesture models than did control condition participants, who never engaged in this activity, $F(2, 33) = 14.32$, $p<0.01$ (see Figure 7). Although demonstrators using gestures produced gesture models no more than demonstrators using both speech and gestures, the two groups differed in the type of information they chose to convey through by models. Whereas demonstrators in the gesture-only condition used gesture modeling to convey information about both structure and action, demonstrators using speech used gestural modeling only to convey information about structure. Speakers either conveyed information about actions within steps with words only, or they entirely omitted this information from their demonstrations.

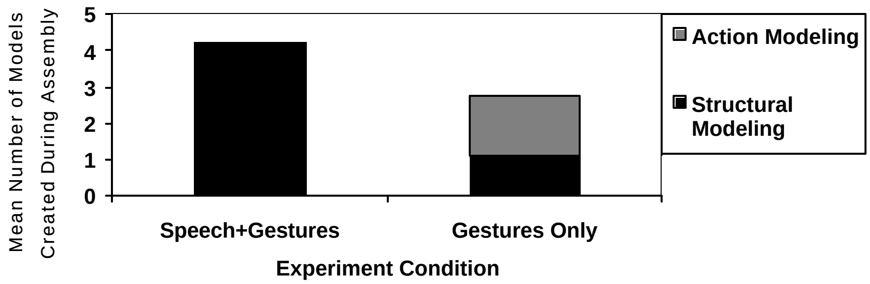

Figure 7 Differences in mean number and type of gestural modeling as a function of experiment condition

How do speech and gesture interact? In the speech and gesture condition, small components and large ones can be referred to by speech or gesture or both. Participants made significantly more verbal references to small object connectors (M=15.42) than to large object parts (M=9.67), t(22)=3.60, p<0.05. This is of particular interest since the reverse pattern was found for these participants' gestures; overall, participants used gestures more frequently to refer to large components than to small ones (Figure 8). This likely occurs because the small parts are hard to see, smaller than a pointing finger for the most part; speech is not constrained by visibility. Also, labels for small parts and connectors, for example, screw and hole, are more specific than labels for the large parts, for example, side and middle boards.

Verbal references were more frequently accompanied by gestures (M=15.08) than in isolation (M=10.00), t(22)=2.33, p<0.05. For large object parts, verbal responses were more frequently accompanied by gestures (M=6.08) than not (M=3.58), t(22)=2.62, p<0.05. However, for small object parts, participants were as likely to make verbal references with gestural accompaniments as to make verbal references that lacked gestural accompaniments, p>0.05 (Figure 9). Gestures almost never occurred without accompanying verbal references, but on occasions where gestures did occur in isolation, more were directed toward small object connectors (M=1.20), than toward large object parts (M=0.083).

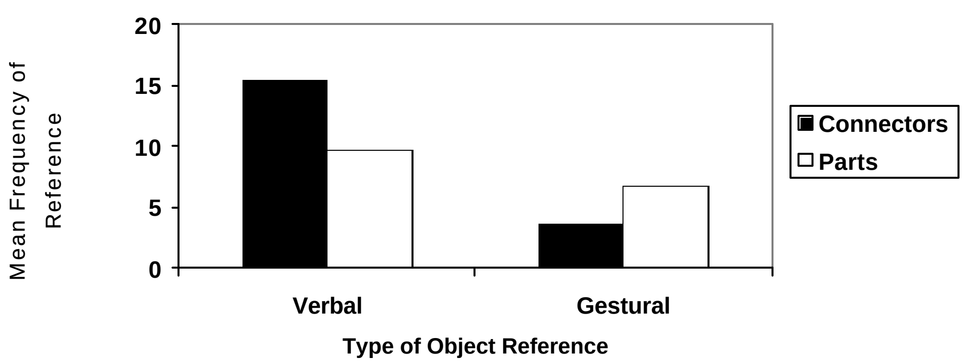

Figure 8 Mean frequency of referencing objects using words vs. gestures in the speech + gesture condition.

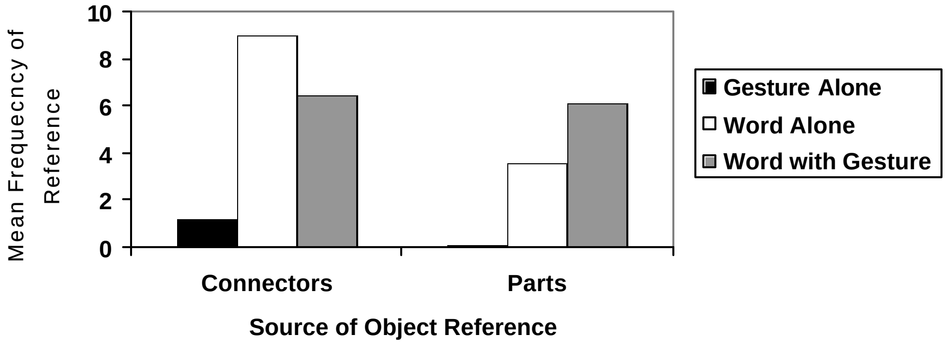

Figure 9 Mean frequencies of word-gesture co-occurrences by function in the speech + gesture condition.

How does spatial ability affect assembly performance? Those higher in spatial ability on the Vandenburg MRT assembled the TV cart faster than those low in spatial ability on the first attempt, $r=-0.41$, $p<0.05$, but not on reassembly ($p>0.05$). There was no correlation between first and second assembly times. In the gesture only condition (but not in the gesture and speech conditions), those higher in spatial ability showed fewer attachment actions to the camera during demonstration $r=-0.63$, $p<0.05$.

4 General Discussion

How do people demonstrate how to do something, using actions, language, and gesture? Here, participants first learned to assemble a TV cart, using a photograph of the completed object as a guide. Then some simply reassembled the cart, but others reassembled in order to instruct others how to do it. Half the communicators were allowed to use speech and gesture, and half were allowed only gesture, with the intention of making a video for non-English speakers. We have examined some of the gestures and speech used, both to understand the ways people use gestures and speech to communicate and to translate these devices to augment animated graphics that instruct.

Turning an ordinarily noncommunicative assembly task into a communicative task adds gestures to the actions used for assembly and also changes the way the assembly is done. Communicators segmented the assembly into clear, discrete steps. Communicators also directed assembly actions so that they would be visible to viewers, and performed these actions in ways that would accommodate viewers' perspective. Making assembly actions visible to the viewer occurred most in the gesture alone condition, where nearly all steps were made visible. The addition of language may allow verbal explanation of assembly actions to compensate for lack of visibility.

Making assembly actions visible to viewers was not the only way that demonstrations differed from reassemblers. Demonstrators supplemented the action assembly with gestures and talk. So far, we have analyzed two prevalent types of gestures, exhibiting, or holding a piece for the viewer, and pointing, using a finger to point to a component. We compared the frequencies of these gestures to small parts and connectors, such as holes and screws, and to large components, such as sideboards. These gestures did not occur at all in the reassembly control condition. As expected, both kinds of gestures were frequent in demonstrations, even more so in the gesture only condition. There were subtle differences in the ways pointing and exhibiting were used to indicate small connectors and large component parts. In addition, the use of these gestures interacted with use of language.

Demonstrators more frequently pointed to small parts and connectors and exhibited large component parts. This appears to be an effect of visibility and mobility. Some parts were too small to be held and seen, such as screws; some connectors, like holes, could not be separated for exhibiting. These naturally received points more than exhibits. The small parts were also more likely to be named. The larger parts that could be separated and were visible to the viewer were exhibited..

We examined a third, less frequent, class of gestures, a sequence of gestures designed to model aspects of the assembly. For the speech and gesture condition, gesture models were frequent, and used primarily to indicate the structure of the TV cart. For the gesture only condition, gesture models indicated the assembly actions needed to assemble the TV cart as well as the structure. Modeling suggests a technique that might improve animations: previewing complex steps before showing them.

The speech and gesture condition allows examining interactions between speech and gesture for small connectors and large components. In general, language was more frequently used to indicate small connectors and gesture to indicate large components, paralleling the use of points and exhibits. Language may be more important for indicating when parts are small than when they are large. In addition, terms for small parts and connectors are more specific, as well as less ambiguous, than terms for the large component parts, which do not have standard names. Consistent with this analysis, indicating connectors was more frequent via language alone and indicating components parts via both language and gestures. Speech replaced pointing more readily than it replaced exhibiting. This suggests another device that might improve animations: enlarging small parts in animations, as they are done in static visualizations such as maps.

Comparing the speech and gesture condition to the gesture only condition suggests that restricting communication to gestures alone may have facilitated communication. Demonstrators free to use both speech and gesture frequently used gestures to accompany speech and also frequently used speech without accompanying gestures, but they almost never used gestures in isolation from accompanying speech. Altogether, speech reduced non-verbal communication devices, both gestures and visibility of actions to viewer. Speakers frequently used gestures to identify objects but rarely used gestures to show how to use those objects. Speaking demonstrators also showed assembly actions to the camera less frequently than nonspeaking demonstrators.

Spatial ability had a limited impact on how people performed in the reassembly phase of the task. Although people low in spatial ability took longer to learn how to build the TV cart the first time, spatial ability level had no effect on the length of time people took to reassemble the TV cart. In other words, once people low in spatial ability had knowledge of the task and had experience doing it, performance-time deficits disappeared. Also, while spatial ability did not predict quantity of gesture usage, it did predict how often demonstrators using gestures alone made their actions visible to the camera; demonstrators with high spatial ability made fewer of their actions visible to the camera. High spatial ability may have made these demonstrators more oblivious to difficulties and uncertainties that unskilled viewers could encounter during assembly.

What are the implications for design of communicative graphics? For this task, assembly of a TV cart, users prefer graphics to text [19]. The techniques used in the gesture only condition, then, are of particular import for design of graphics, especially animated graphics. Several lessons can be drawn. Segment the task into steps. Make the assembly of each step visible to viewers. Preview complex steps. Change spatial and temporal scale when needed. Use extra-pictorial devices. Indeed, these are also in accord with users' preferences [19]. The gestures suggest specific pictorial devices. Points are similar to arrows. One function of each is to indicate; that is the function studied here. Exhibiting is analogous to drawings of individual components.

Typically, animations do not segment actions into meaningful steps. Animations often do not supplement the visual information with gestures that indicate, exhibit, and explain. Animations do not normally change spatial or temporal scale. These demonstrations give insight into why animated graphics fail to benefit learning more than static ones. For animations to be effective, they must go beyond merely showing a process; to be effective animations must also explain how a process works. As for all visualizations, effective animations are not necessarily those that present more information more realistically; clarity, quality, and coherence of informational delivery likely play important roles in effectiveness of animations. It remains to be seen whether adding these insights from communicative demonstrations to animations will augment their effectiveness.

Acknowledgments

This research was supported in part by ONR grants NOOO14-PP-1-O649, N000140110717, and N000140210534 to Stanford University. We are grateful to Herb Clark for his insights and advice.

References

[1] Kaiser, M. K., Proffitt, D.R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 669-690.

[2] Rieber, L. P. (1991). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology, 83*, 318-328.

[3] Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57*, 247-262.

[4] Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering and communicating structure in events. *Journal of Experimental Psychology: General*, 136, 29-58.

[5] Tversky, B. (1995). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.). *Understanding images*. (pp.29-53). New York: Springer-Verlag.

[6] Tversky, B. (2000) Some ways that maps and graphs communicate. In Freksa, C., Brauer, W., Habel, C and Wender, K. F. (Eds.), *Spatial cognition II: Integrating abstract theories, empirical studies, formal methods, and practical applications.* Pp. 72-79. N. Y.: Springer.
[7] Tversky, B, Zacks, J., Lee, P. U., & Heiser, J. (2000). Lines, blobs, crosses, and arrows: Diagrammatic communication with schematic figures. In M. Anderson, P. Cheng, and V. Haarslev (Editors). *Theory and application of diagrams*. Pp. 221-230. Berlin: Springer.
[8] Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
[9] Tversky, B., & Lee, P. U. (1998). How space structures language. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition: An interdisciplinary approach to representation and processing of spatial knowledge*. (pp. 157-175). Berlin: Springer-Verlag.
[10] Tversky, B., & Lee, P. U. (1999). Pictorial and verbal tools for conveying routes. In Freksa, C., & Mark, D. M. (Eds.). *Spatial information theory: cognitive and computational foundations of geographic information science*. (pp. 51-64). Berlin: Springer.
[11] Heiser, J. and Tversky, B. (submitted). Mental models of complex systems: Structure and function.
[12] Rauscher, F. H., Krauss, R. M, & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, 7, 226-231.
[13] Cohen, A. A., & Harrison, R. P. (1973). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology, 28,* 276-279.
[14] Gullberg, M. (1998). Gesture as a communication strategy in second-language discourse: A study of learners of French and Swedish. Lund: Lund University Press.
[15] Cassell, J., McNeill, D. and McCullough, K.-E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics and Cognition*, 7. 1-33.
[16] Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communication, 27,* 54-63.
[17] Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61,* 743-754.
[18] Vandenburg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual & Motor Skills*, 47, 599-604.
[19] Heiser, J., Tversky, B., Agrawala, M., & Hanrahan, P. (2003). Cognitive design principles for visualizations: Revealing and instantiating. *Cognitive Science.* Manuscript submitted for publication.
[20] Emmorey, K., Tversky, B., & Taylor, H. A. (2000) Using space to describe space: Perspective in speech, sign, and gesture. *Journal of Spatial Cognition and Computation,* 2, 157-180.